UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 10/698,629 | 10/31/2003 | Robert W. Bossemeyer | 1048.002US1 | 3387 |

7590 07/20/2007

Law Offices of Michael Dryja
704 228th Avenue NE PMB 694
Sammamish, WA 98074

| EXAMINER | |
|---|---|
| LEE, GINA W | |
| **ART UNIT** | **PAPER NUMBER** |
| 2609 | |

| MAIL DATE | DELIVERY MODE |
|---|---|
| 07/20/2007 | PAPER |

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

| Office Action Summary | Application No. | Applicant(s) | |
|---|---|---|---|
| | 10/698,629 | BOSSEMEYER ET AL. | |
| | Examiner | Art Unit | |
| | Gina W. Lee | 2609 | |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

1)☒ Responsive to communication(s) filed on *31 October 2003*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

4)☒ Claim(s) *1-20* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-20* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

**Application Papers**

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *31 October 2003* is/are: a)☐ accepted or b)☒ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

## DETAILED ACTION

### *Drawings*

1. The drawings are objected to under 37 CFR 1.83(a) because they contradict the specification. Drawings that are labeled frequency-time are described in the specification as depicting the change in amplitude over time. (See particularly Figures 3, 4, 6A, 7, 8D, 8E, 9D, and 9E.) Any detail that is essential for a proper understanding of the disclosed invention should be shown in the drawing. MPEP § 608.02(d). Corrected drawing sheets in compliance with 37 CFR 1.121(d) are required in reply to the Office action to avoid abandonment of the application. Any amended replacement drawing sheet should include all of the figures appearing on the immediate prior version of the sheet, even if only one figure is being amended. The figure or figure number of an amended drawing should not be labeled as "amended." If a drawing figure is to be canceled, the appropriate figure must be removed from the replacement sheet, and where necessary, the remaining figures must be renumbered and appropriate changes made to the brief description of the several views of the drawings for consistency. Additional replacement sheets may be necessary to show the renumbering of the remaining figures. Each drawing sheet submitted after the filing date of an application must be labeled in the top margin as either "Replacement Sheet" or "New Sheet" pursuant to 37 CFR 1.121(d). If the changes are not accepted by the examiner, the applicant will be notified and informed of any required corrective action in the next Office action. The objection to the drawings will not be held in abeyance.

## *Claim Rejections - 35 USC § 103*

2. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

3. Claims 1-8 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kametani (US 5,091,948) in view of Prezas (US 4,561,102).

4. Regarding independent **claim 1**, Kametani teaches a method, comprising:

receiving a signal representing digitized, sampled human speech (column 2, lines 42-50, A/D converter digitizes the analog waveform at a particular sampling rate);

locating at least one speech segment within the signal (Fig 2, column 2, line 51-column 3, line 5, voiced speech is located); and

locating one or more higher energy sections within each speech segment within the signal (Fig 2, column 2, line 51-column 3, line 5, voiced speech is located). However, Kametani does not teach the method of pitch detection.

Prezas teaches a method comprising:

locating a plurality of glottal events within each speech segment within the signal, (Fig. 4, column 4, lines 50-58, pulses are measured and their locations recorded); and,

confirming the plurality of glottal events located within each speech segment within the signal, including registering each of at least one of the plurality of glottal events with adjacent glottal events (Fig. 4, column 6, line 24-column 7, line 14, pulses are placed in order of occurrence and the interval between them is measured and recorded).

Pitch detection is a very useful procedure that is used in all kinds of speech signal processing, including speaker verification. Prezas's method is of especial value because pitch detection in time domain LPC analysis of speech is effected in real time. As Kametani teaches a generic system for speaker verification that may use pitch as one of the identifying speech features, and Prezas teaches an accurate method of pitch detection, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Kametani's method with Prezas's.

5. With respect to dependent **claim 2**, Prezas teaches a method of receiving a signal representing digitized, sampled human speech,

comprising sampling the human speech to digitize the human speech, yielding the signal (Fig. 1, column 4, lines 8-9, coder (11) samples speech and digitizes the sample amplitudes).

While Prezas does not explicitly state so, it would have been obvious to one of ordinary skill in the art that the speech source (10) may also be speech that as been previously recorded, as the nature of the speech source does not affect Prezas's invention.

6. With respect to dependent **claims 3 and 4**, Kametani teaches a method of locating a speech segment within the signal comprising:

determining a start point and an end point of each speech segment by determining an energy within the signal and examining the energy for regions above a threshold (Fig 2, column 2, line 51-column 3, line 5, voiced speech is located).

7. With respect to dependent **claim 5**, Kametani teaches a method of locating a higher energy section comprising:

determining regions within each speech segment where an energy is at least a percentage of a peak energy with in the speech segment (Fig 2, column 2, line 51-column 3, line 5, voiced speech is located).

8. With respect to dependent **claim 6**, Prezas teaches a method of locating a plurality of glottal events comprising:

subjecting each higher energy section within the speech segment to a linear predictive coefficient (LPC) analysis, yielding a LPC residual error signal for each higher energy section (column 4, lines 15-33, LPC analysis is performed and a residual signal is created);

locating a number of largest peaks within the LPC residual error signal for each higher energy section that have a minimum separation between adjacent of the peaks (column 4, line 64-column 5, line 3, amplitudes of the peaks are measured in order of decreasing height and stored in the pulse amplitude storage (48). Locations are stored in the pulse location storage (49)); and

locating the plurality of glottal events within the speech segment as corresponding to the number of largest peaks within the LPC residual error signal that have the minimum separation (column 5, lines 13-18, pulses that are closer than a minimum separation are eliminated from consideration).

9. With respect to dependent **claim 7**, Prezas teaches a method of subjecting a higher energy section to LPC analysis, yielding the LPC residual error signal, comprising:

determining the LPC residual error signal as the square of the difference between the higher energy section and an LPC-derived model of the higher energy section (column 4, lines 10-13, samples of the LPC model are combined with each actual sample to find the best least square fit).

10. With respect to dependent **claim 8**, Prezas teaches a method of locating the number of largest peaks within the LPC residual error signal that have the minimum separation between adjacent of the peaks comprising:

from all the largest peaks within the LPC residual error signal, removing those peaks that lack the minimum separation between adjacent of the peaks (column 5, lines 13-18, pulses that are closer than a minimum separation are eliminated from consideration).

11. Claims 9 and 14-16 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kametani (US 5,091,948) in view of as applied to claim Prezas (US 4,561,102) above, and further in view of Petrushin ("Pitch-Synchronous Speech Signal Segmentation and Its Applications").

12. Regarding dependent **claim 9**, Kametani and Prezas together teach the method of claim 1, which claim 9 is dependent upon, in its entirety (as outlined above) but are silent as to the comparison of actual glottal events or adjustments to them.

Petrushin discusses a method of speech signal segmentation comprising:

comparing a first glottal event and a second glottal event of the adjacent pair of glottal events to determine a pair-wise distance between the first and the second glottal events (pages 322-323, adjacent speech segments are compared and the distance between them is measured); and

adjusting boundaries of at least one of the first glottal event and the second glottal event to minimize the pair-wise distance between the first and the second glottal events, maximizing similarity of the first and the second glottal events of the adjacent pair (pages 325-326, the endpoints of the quasi-periodic segment are adjusted to decrease the error between them).

Kametani and Prezas together teach a method of locating glottal events, but Prezas, while allowing a margin of error in the pulse location calculation, is silent as to a method of adjusting calculations to increase the similarity between sections of the signal. Petrushin teaches a method by which the segmentation of the waveform (which corresponds to glottal pulse locations) may be adjusted improve the accuracy. The sections of the signal are compared and changing the endpoints of the sections minimizes the error between them. Conversely, Petrushin suggests using a method of segmentation of the signal by glottal pulse positions (like Prezas) but does not recommend any particular system of speaker verification to implement the method. It is understood that it would need to be part of a larger speaker recognition system to be useful, such as

Kametani's, since it can be assumed that it would be necessary to obtain and store the signal that Petrushin's method analyzes. Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Kametani's method of speaker recognition with Petrushin's method of comparing and adjusting glottal events.

13. With respect to dependent **claim 14**, this claim is very similar to claims 6 and 9, and is rejected for the same reasons.

14. With respect to dependent **claim 15**, this claim is very similar to claim 9, and is rejected for the same reasons.

15. With respect to dependent **claim 16**, this claim is very similar to claim 9, and is rejected for the same reasons.

16. Claims 10-13 are rejected under 35 U.S.C. 103(a) as being unpatentable over Petrushin ("Pitch-Synchronous Speech Signal Segmentation and Its Applications").

17. Regarding independent **claim 10**, Petrushin teaches a method comprising:

for each adjacent pair of glottal events within each of a plurality of speech segments within a signal representing digitized, sampled human speech,

comparing a first glottal event and a second glottal event of the adjacent pair of glottal events to determine a pair-wise distance between the first and the second glottal

events (pages 322-323, adjacent speech segments are compared and the distance between them is measured); and,

adjusting boundaries of at least one of the first glottal event and the second glottal event to minimize the pair-wise distance between the first and the second glottal events (pages 325-326, the endpoints of the quasi-periodic segments are adjusted to decrease the error between them).

18. Although Petrushin does not explicitly state implementing the method using a computer, it would have been obvious to one of ordinary skill in the art that digital signal processing requires a computer, as well as a computer readable medium upon which to store the digitized signal being analyzed.

19. With respect to dependent **claim 11**, Petrushin teaches

adjusting the boundaries of at least one of the first glottal event and the second glottal event comprises adjusting at least one of a start point and an end point of at least one of the first glottal event and the second glottal event (pages 325-326, the endpoints of the quasi-periodic segment are adjusted to decrease the error between them).

20. With respect to dependent **claim 12**, Petrushin teaches

adjusting the boundaries of at least one of the first glottal event and the second glottal event maximizes similarity of the first and the second glottal events (pages 325-326, the endpoints of the quasi-periodic segment are adjusted to decrease the error between them).

21. With respect to dependent **claim 13**, Petrushin teaches

initially locating a plurality of glottal events within each speech segment within the signal (page 321, initial segmentation may be done by location of glottal pulse positions).

22. Claims 17-20 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kametani (US 5,091,948) in view of as applied to claim Prezas (US 4,561,102) above, and further in view of Radová et al. ("An Approach to Speaker Identification Using Multiple Classifiers").

23. Regarding independent **claim 17**, Kametani teaches a speaker verification system comprising:

a computer-readable medium having stored thereon a plurality of events extracted from human speech (Fig 1, column 5, lines 17-23, pattern memory (43) containing reference parameters); and

a recording device to record further human speech and store a signal representing the further human speech on the computer-readable medium (Fig 1, column 2, lines 38-50 and column 3, microphone (1) receives human speech and provides the digitized signal to a buffer (10) in memory)

but does not teach any method of generating glottal events.

Prezas teaches a method:

to generate a plurality of second glottal events from the signal (Fig. 4, column 4, lines 50-58, pulses are measured and their locations recorded); and

to confirm the plurality of second glottal events by registering each second glottal event with adjacent second glottal events (Fig. 4, column 6, line 24-column 7, line 14, pulses are placed in order of occurrence and the interval between them is measured and recorded)

but does not teach any method of comparing the glottal events to each other.

Radová teaches a method of speaker verification

to compare the plurality of second glottal events with the plurality of first glottal events to determine whether the further human speech recorded matches the previously recorded human speech (pages 1136-1137, a test speech segment and a reference speech segment are compared and classified in order to identify the speaker of the test segment).

Kametani and Prezas together teach a method of locating glottal events, but they do not teach any method of using the glottal event as an feature to perform speaker verification. Kametani teaches a method of estimating the glottal source and the glottal pulse in order to perform speaker verification. Radová teaches a method by which the actual measured and digitized waveform itself can be used in speaker verification. As computing speed and capability increase, it has become feasible to use the features of the actual waveform itself for more accurate identification. On the other hand, Radová implies that the speech segment is part of a larger speech signal, and has been isolated using some method, such as Prezas's. Radová does not recommend any particular system of speaker verification to implement her method, but it is understood that it would need to be part of a larger speaker recognition system to be useful, such as Kametani's. Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Kametani's method of speaker recognition with Prezas's method of locating glottal events and Radová's method of comparing glottal events for speaker verification.

24. With respect to dependent **claim 18**, Prezas teaches

confirming the plurality of second glottal events by registering each second glottal event with adjacent second glottal events (Fig. 4, column 6, line 24-column 7, line 14, pulses are placed in order of occurrence and the interval between them is measured and recorded).

25. With respect to dependent **claim 19**, Prezas teaches

a computer program stored on the computer-readable medium (column 5, lines 38-509, column 9, line 62-column 10, line 4, column 10, lines 40-51, Prezas's method can be implemented by a system containing memory, processors, and programmable logic).

26. Regarding independent **claim 20**, Kametani teaches a system comprising

means for recording human speech and for storing a signal representing the human speech on a computer-readable medium (Fig 1, column 2, lines 38-50 and column 3, microphone (1) receives human speech and provides the digitized signal to a buffer (10) in memory) having previously stored thereon a plurality of first glottal events extracted from previously recorded human speech (Fig 1, column 5, lines 17-23, pattern memory (43) containing reference parameters)

but does not teach any method of locating glottal events.

Prezas teaches a method of:

generating a plurality of second glottal events from the signal (Fig. 4, column 4, lines 50-58, pulses are measured and their locations recorded), for confirming the plurality of second glottal events by registering each second glottal event with adjacent second glottal events (Fig. 4, column 6, line 24-column 7, line 14, pulses are placed in order of occurrence and the interval between them is measured and recorded)

but does not teach any method of comparing the identified glottal events with each other.

Radová teaches a method of speaker verification for

comparing the plurality of second glottal events with the plurality of first glottal events to determine whether the further human speech recorded matches the previously

recorded human speech (pages 1136-1137, a test speech segment and a reference speech segment are compared and classified in order to identify the speaker of the test segment).

Kametani's system, modified by the methods of Prezas and Radová, teach the claimed system in its entirety.

***Conclusion***

27. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

French (US 3,466,394) discloses a voice verification system.

Lindenberg (US 3,940,565) discloses a time domain voice verification system.

Doddington et al. (US 3,700,815) discloses a speech verification system by time alignment of acoustic parameters.

Sakoe (US 4,403,114) discloses a speaker recognition system by pattern matching of time-normalized signals.

Ma et al. (US 6,470,308) discloses a system for detecting instants of glottal closure.

Radová et al. ("Method for the Segmentation of Voiced Speech Signals Into Pitch Period Segments") discloses a method of segmenting the speech signal into glottal events.

LaRoy Berg et al. ("Investigating Speaker Features from Very Short Speech Records") discloses a method of speaker verification using sections of glottal signals.

28. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Gina W. Lee whose telephone number is (571) 270-3139. The examiner can normally be reached on Monday to Thursday, 6:30 AM - 5:00 PM EST.

29. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Alexander Eisen can be reached on (571) 272-2687. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

30. Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

Alexander Eisen
SPE
Art Unit 2609

GWL